
**Caracterización de los niveles de
contaminación de cuatro zonas de
Bogotá por análisis de datos tipo
panel**



Presentado por

Jessica Ximena Moreno

Wilson Sandoval Rodriguez

LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Fundación Universitaria Los Libertadores

Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2019

**Caracterización de los niveles de
contaminación de cuatro zonas de
Bogotá por análisis de datos tipo
panel**



LOS LIBERTADORES

Presentado por
Jessica Ximena Moreno Cortes
Wilson Sandoval Rodriguez

en cumplimiento parcial de los requerimientos para optar al título
de

Especialista en Estadística Aplicada

Quien asesora
Heivar Yesid Rodríguez Pinzón
Profesor

Fundación Universitaria Los Libertadores
Facultad de Ingeniería y Ciencias Básicas
Especialización en Estadística Aplicada
Bogotá D.C, Colombia
2019

Notas de aceptación



LOS LIBERTADORES

FUNDACIÓN UNIVERSITARIA

Firma del presidente del jurado

Firma del jurado

Firma del jurado

Bogotá D.C., Junio de 2019.



LOS LIBERTADORES

FUNDACIÓN UNIVERSITARIA

Las directivas de la Fundación Universitaria Los Libertadores, los jurados calificadores y el cuerpo docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores y a los resultados de su trabajo.

Índice general

1	Introducción	3
2	Planteamiento del Problema	5
2.1	Objetivos	8
2.1.1	Objetivo General	8
2.1.2	Objetivos Específicos	8
3	Marco Normativo	9
3.1	Calidad del aire.	9
3.1.1	Material Particulado (PM)	11
3.1.2	Gases contaminantes	11
4	Marco Metodológico	13
4.1	Fuentes de información	13
5	Análisis y Resultados	15
5.1	Técnica de análisis de datos	15
5.1.1	Datos de Panel	15
5.1.2	Especificación general de un modelo de datos de panel	15
5.1.3	Componentes del error	16
5.1.4	Ventajas y Desventajas de la Técnica de Datos de Panel	16
5.2	Resultados	18
5.2.1	Panel de datos Dióxido de Azufre	19
5.2.2	Panel de datos Ozono	28
5.3	Análisis Multivariado	37
6	Conclusiones y Recomendaciones	43

Índice de figuras

5-1	Secciones transversales combinadas	20
5-2	Secciones Transversales apiladas	21
5-3	Prueba de Normalidad	24
5-4	Gráfico de estimación	26
5-5	Secciones transversales O_3	29
5-6	Secciones transversales apiladas O_3	29
5-7	Prueba de normalidad	32
5-8	Gráfico de estimación	34
5-9	Índice de la calidad del aire	37
5-10	$PM_{2.5}$ por individuo	38
5-11	PM_{10} por individuo	38
5-12	SO_2 por individuo	38
5-13	MCA Categorías	39
5-14	MCA categorías con individuos	40
5-15	MCA variables	41
5-16	$mca \cos^2$	42

Índice de tablas

3-1	Niveles a partir de 2018	10
3-2	Niveles máximos permisibles de contaminantes en el aire para el año 2030	10
4-1	Ubicación estaciones RMCAB	14
5-1	Tabla Datos panel	16
5-2	Estadísticas Descriptivas	19
5-3	Prueba igualdad de medias.	21

5-4	Prueba igualdad de varianzas.	22
5-5	Estimación del modelo	22
5-6	R^2 y Durbin -Watson	23
5-7	Add your caption here.	24
5-8	Prueba dependencia de los residuales	25
5-9	Heterocedasticidad de los residuales	25
5-10	Cross-section Fixed effects SO_2	26
5-11	Period Fixed Effects	27
5-12	Estadísticas descriptivas ozono	28
5-13	Igualdad de medias.	30
5-14	Igualdad de varianzas Ozono	30
5-15	Add your caption here.	31
5-16	Add your caption here.	33
5-17	Prueba dependencia residuales.	33
5-18	Prueba heterocedasticidad	34
5-19	Cross-Section Fixed effects O_3	35
5-20	Period Fixed Effects.	35

Caracterización de los niveles de contaminación de cuatro zonas de Bogotá por análisis de datos tipo panel

Resumen

Varias organizaciones se encuentran realizando colaboraciones para contrarrestar los problemas ambientales generados por la producción de emisiones contaminantes, generados por el constante desarrollo y modernización urbanística, en las principales fuentes de contaminación del aire se encuentran los productos derivados del petróleo y por fuentes naturales, donde se producen sustancias contaminantes que generan en la salud humana graves enfermedades como cerebrovasculares, cánceres de pulmón y neuropatías crónicas y agudas.

El objetivo de este estudio es proponer un modelo de datos panel de efectos fijos para determinar los factores más significativos que influyen en la concentración de gases contaminantes atmosféricos en cuatro zonas de Bogotá. De acuerdo con este fin, la pregunta de investigación es: ¿Cuáles son los factores más importantes sobre la concentración de los gases contaminantes atmosféricos que influyen en el índice de contaminación del aire en cuatro zonas de Bogotá?

La pregunta de investigación se responde a través de un modelo de datos panel de efectos fijos para determinar los factores más significativos que influyen en la concentración de gases contaminantes atmosféricos, mediante un análisis descriptivo multivariado entre los contaminantes atmosféricos SO_2 , NO_2 , CO , O_3 , material particulado y algunas variables meteorológicas.

Palabras claves: gases contaminantes, Material Particulado, Datos Panel,

Capítulo 1

Introducción

Actualmente las grandes ciudades se enfrentan a problemas ambientales generados por la producción de emisiones contaminantes, el constante desarrollo y modernización urbanística hace que las personas requieran en sus actividades diarias el uso de energías no renovables y uno de los principales retos ambientales son los niveles de concentración de emisiones de contaminantes en el aire, los cuales son nocivos para la salud humana y la naturaleza en general.

Dentro de las principales fuentes de contaminación del aire que se encuentran en el diario vivir son aquellos productos derivados del petróleo como el combustible para vehículos de servicio público y particular, las industrias, los hogares y adicionalmente aquellas emisiones que son producidas por fuentes naturales. Estas fuentes producen sustancias contaminantes como O_3 , CO , SO_2 , NO_2 , $PM_{2.5}$ y PM_{10} , que generan en la salud humana problemas graves como enfermedades de cáncer, problemas respiratorios, cardiovasculares entre otros, según estudios de la Organización Mundial de la Salud [O.M.S. \(2018\)](#), que en los últimos años ha establecido manuales y protocolos para dar lineamientos a los países en cuanto a los niveles permisibles de los contaminantes atmosféricos.

El notable deterioro de la calidad del aire en Bogotá y su directa relación con problemas de salud respiratoria y cardiaca ha generado una creciente preocupación por parte de autoridades ambientales y de salud pública de la ciudad. A pesar de los esfuerzos realizados en años recientes, el problema de contaminación atmosférica es cada vez más severo. La ciudad de Bogotá cuenta con la Red de Monitoreo de Calidad del Aire de Bogotá denominada RMCAB para el año 2016 RMCAB contaba con trece estaciones las cuales doce de ella recolectaban información sobre la concentración de contaminantes y variables meteorológicas, y la estación Bolivia encargada de hacer recolección de información únicamente de las variables meteorológicas, las estaciones se encuentran ubicadas de manera estratégica con el fin de obtenga la mayor cobertura de la ciudad, esta información es

procesada en la estación principal en donde se valida y se evalúa a partir de los estándares de calidad del aire en Bogotá

En el presente trabajo se pretende determinar cuales son los factores más importantes sobre la concentración de los gases contaminantes atmosféricos (SO_2 , NO_2 , CO , O_3) que influyen en el índice de contaminación del aire en cuatro zonas de la ciudad de Bogotá

Los datos utilizados en el desarrollo de este trabajo tienen como fuente los datos publicados de las estaciones de la Red de Monitoreo de Calidad del Aire de Bogotá - RMCAB, que se obtienen a través de la página web de la secretaria distrital de ambiente y fueron tratados mediante análisis de datos tipo panel esta técnica permite capturar la heterogeneidad no observable ya sea entre unidades individuales de estudio como en el tiempo, los datos en panel suponen e incorporan, en el análisis los individuos las zonas de Bogotá consideradas para el estudio de este trabajo. Los análisis de series de tiempo y de corte transversal no tratan de controlar esta heterogeneidad corriendo el riesgo de obtener resultados sesgados.

Siendo el objetivo del trabajo proponer un modelo de datos panel de efectos fijos para determinar los factores más significativos que influyen en la concentración gases contaminantes atmosféricos (SO_2 , NO_2 , CO , O_3) en cuatro zonas de Bogotá.

Capítulo 2

Planteamiento del Problema

Actualmente las grandes ciudades se enfrentan a problemas ambientales generados por la producción de emisiones contaminantes, el constante desarrollo y modernización urbanística hace que las personas requieran en sus actividades diarias el uso de energías no renovables y uno de los principales retos ambientales son los niveles de concentración de emisiones de contaminantes en el aire, los cuales son nocivos para la salud humana y la naturaleza en general.

Dentro de las principales fuentes de contaminación del aire que se encuentran en el diario vivir son aquellos productos derivados del petróleo como el combustible para vehículos de servicio público y particular, las industrias, los hogares y adicionalmente aquellas emisiones que son producidas por fuentes naturales. Estas fuentes producen sustancias contaminantes como O_3 , CO , SO_2 , NO_2 , $PM_{2.5}$ y PM_{10} . Estas sustancias contaminantes generan en la salud humana problemas graves como enfermedades de cáncer, problemas respiratorios, cardiovasculares entre otros.

En los últimos años la Organización Mundial de la Salud a aumentado los estudios sobre la calidad del aire [“WHO — Ambient and household air pollution and health” \(2018\)](#) donde se han establecido manuales y protocolos para dar lineamientos a los países en cuanto a los límites permisibles de los contaminantes atmosféricos, con el fin de producir mecanismos políticos que permitan reducir las emisiones contaminantes en el aire. En Colombia, actualmente a través del Ministerio de Ambiente y Desarrollo Sostenible [Abel y Flórez \(2018\)](#), se obtuvo la resolución 2254 de 1 de noviembre del 2017 el cual establece la norma de calidad del aire para todo el territorio nacional, con el fin de garantizar para el año 2030 niveles de contaminantes atmosféricos bajos que permiten tener un ambiente sano y reducir los riesgos de salud para la población.

La contaminación del aire es la principal preocupación para la Organización Mundial

de la Salud, la cual en 2006 divulgó la “Guías de calidad del aire de la OMS relativas al material particulado, el ozono, el dióxido de nitrógeno y el dióxido de azufre” *OPS OMS — Calidad del aire — Productos científicos y técnicos* (s.f.). Esta guía establece criterio de calidad del aire, con el fin de poder reducir las emisiones de las sustancias contaminantes y sus efectos en el aire que afectan la salud de los seres vivos. La guía fue publicada por primera vez en el año 1987 y fue actualizada en el año 1997, en donde se basaron en pruebas que realizaron expertos científicos del aquel momento. Una segunda edición fue publicada en *Air quality Guidelines for Europa*, donde se tuvieron en cuenta las investigaciones realizadas en donde se determinó que los países de ingreso bajo y mediano presentaron altos niveles de contaminación del aire, alcanzando un máximo.

La Organización Mundial de la Salud realizó un nuevo estudio teniendo en cuenta los resultados científicos anteriores, por lo cual generó la guía de calidad del aire. Esta guía tiene como principal objetivo establecer políticas y dar a conocer diversas opciones políticas para la gestión de la calidad de aire en los diferentes partes del mundo. En esta guía de calidad del aire contiene información común sobre las cuatro sustancias contaminantes material particulado (MP), ozono (O₃), dióxido de nitrógeno (NO₂) y dióxido de azufre (SO₂).

Una nueva actualización de esta guía de calidad del aire de la Organización Mundial de la Salud se realizó en octubre de 2006, donde se seleccionó un grupo de expertos en toxicología, epidemiología evaluación de exposición a la calidad del aire, manejo de la calidad del aire y políticas públicas. Con nuevas evidencias sobre la contaminación del aire, los expertos sugirieron establecer metas de mediano plazo, con el fin de poder reducir las emisiones contaminantes y mejorar la salud pública en aquellas ciudades con altos niveles de contaminación. [O.M.S. \(2018\)](#)

En Colombia en los últimos años ha aumentado la preocupación por los impactos de la contaminación atmosférica en la salud y del medio ambiente. En el año 1967 se instalaron las primeras redes para vigilar y monitorear la calidad del aire, luego en 1973 se expidió la Ley 23, cuyo objetivo consiste en prevenir y controlar la contaminación del medio ambiente y buscar el mejoramiento, conservación y restauración de los recursos naturales renovables, para defender la salud y el bienestar de todos los habitantes del territorio nacional. A partir de esta ley el Gobierno Nacional expidió en el Código Nacional de Recursos Naturales Renovables y de Protección al Medio Ambiente – Decreto-Ley 2811 de 1974. En 1979, el Congreso de la República aprobó la Ley 9, expidió el Código Sanitario Nacional, la cual se definieron normas, programas y medidas para la protección del medio ambiente. Se facultó al Ministerio de Salud, hoy Ministerio de la Protección Social, para proferir normas para el control de la contaminación atmosférica.

En 1982 se emitió la norma que reguló la emisión y concentración de contaminantes de la atmósfera, también se estableció estándares de calidad del aire y de emisión por fuentes fijas mediante el Decreto 02. Posteriormente, en 2001 se expidió una norma específica para el Distrito Capital mediante la cual se estableció estándares más rigurosos y se incluyeron contaminantes que no habían sido contemplados en la regulación nacional. En 1991 por medio de la Constitución Nacional se estableció una serie de derechos y obligaciones relacionados con el medio ambiente. Para el año 1993 se expidió la Ley 99 por la cual se crea el Ministerio del Medio Ambiente, se reordena el Sector Público encargado de la gestión y conservación del medio ambiente y los recursos naturales renovables y se organiza el Sistema Nacional Ambiental (SINA). En marzo de 2005, el Consejo Nacional de Política Económica y Social aprobó el documento Conpes 3344 que contiene los lineamientos para la formulación de la Política de Prevención y Control de la Contaminación del Aire.

En el año 2008 se realizó una encuesta para el Ministerio de Ambiente, Vivienda y Desarrollo Territorial, en donde se evidencio que el 67% de los encuestados perciben la contaminación como un problema ambiental y 52% consideran la contaminación del aire como el principal problema (Ministerio de Ambiente, 2018). En 2015 por medio de la Resolución 2410, El distrito Capital de Bogotá estableció un indicador multipropósito adimensional llamado Índice Bogotano de Calidad de Aire, IBOCA. Este indicador permite calcular las concentraciones de contaminantes atmosféricos en un momento y lugar de la ciudad. También funciona como indicador de riesgo ambiental por contaminación atmosférica en el marco del Sistema Distrital de Alertas del Sistema Distrital de Gestión del Riesgo y Cambio Climático [Alcaldía Mayor de Bogotá D.C. \(s.f.\)](#) .

En la ciudad de Bogotá, se mide la calidad del aire a través de la Red de Monitoreo de Calidad del Aire (RMCAB), esta red está conformada por 13 estaciones ubicadas en zonas importantes de la ciudad de Bogotá y una red móvil, de tal manera que se puede realizar el seguimiento de las concentraciones de material particulado (PM_{10} , $PM_{2.5}$, PST), de gases contaminantes (SO_2 , NO_2 , CO, O_3) y los parámetros meteorológicos de precipitación, viento, temperatura superficial, radiación solar y humedad relativa, en el último año en la ciudad de Bogotá, la Alcaldía en ocasiones ha declarado alerta amarilla y se han tomado medidas para mitigar los altos niveles de contaminación, el comportamiento de la calidad del aire en Bogotá es evaluado mediante el índice bogotano de calidad del aire (IBOCA) determinando el riesgo ambiental por contaminación atmosférica Y el estado de la calidad del aire de Bogotá.

Los entes reguladores están en condiciones de determinar las afectaciones y brindar recomendaciones en salud y las medidas voluntarias para que la ciudadanía contribuya a

mantener o mejorar la calidad del aire de la ciudad y el indicador IBOCA también funciona como indicador de riesgo ambiental por contaminación atmosférica en el marco del Sistema Distrital de Alertas, del Sistema Distrital de Gestión del Riesgo y Cambio Climático.

Lo cual origina la siguiente pregunta de investigación: **¿Cuáles son los factores más importantes sobre la concentración de los gases contaminantes atmosféricos (SO_2 , NO_2 , CO , O_3) que influyen en el índice de contaminación del aire en cuatro zonas de Bogotá?**

De acuerdo a esta pregunta de investigación se define el siguiente objetivo general y los correspondientes objetivos específicos:

2.1. Objetivos

2.1.1. Objetivo General

Proponer un modelo de datos panel de efectos fijos para determinar los factores más significativos que influyen en la concentración gases contaminantes atmosféricos (SO_2 , NO_2 , CO , O_3) en cuatro zonas de Bogotá.

2.1.2. Objetivos Específicos

- Identificar las relaciones mediante un análisis descriptivo multivariado entre los contaminantes atmosféricos (SO_2 , NO_2 , CO , O_3), material particulado y algunas variables meteorológicas.
- Determinar un modelo de datos panel para la concentración de gases contaminantes de los gases contaminantes atmosféricos (SO_2 , O_3) en cuatro zonas de Bogotá.
- Interpretar los factores más significativos que influyen en la concentración gases contaminantes atmosféricos (SO_2 , O_3) en cuatro zonas de Bogotá.

Capítulo 3

Marco Normativo

3.1. Calidad del aire.

El Índice de calidad del aire (*ICA*) permite comparar los niveles de contaminación del aire de las estaciones de monitoreo que conforman un Sistema de Vigilancia de Calidad del Aire, en un tiempo t que corresponde al período de exposición previsto en la norma para cada uno de los contaminantes que se está midiendo.

Teniendo en cuenta que los contaminantes son monitoreados en todo el país y el índice de calidad del aire se calcula de manera independiente para cada uno de los seis principales contaminantes, los cuales son: PM_{10} , $PM_{2,5}$, SO_2 , NO_2 , O_3 y CO , dicho indicador se calcula a partir de las mediciones de concentración de los contaminantes monitoreados en los Sistemas de Vigilancia de la Calidad del Aire (SVCA) y los puntos de corte preestablecidos para cada uno de los contaminantes, de acuerdo con la siguiente expresión:

$$ICA_p = \frac{I_{Hi} - I_{L_0}}{BP_{Hi} - BP_{L_0}}(C_{pjt} - BP_{L_0}) + I_{L_0} \quad (3-0)$$

- ICA_p := Es el índice de calidad del aire para el contaminante p .
- C_{pjt} := Es la concentración del contaminante medida en la estación de monitoreo de la calidad del aire durante un periodo de tiempo.
- BP_{Hi} := Es el punto de corte mayor o igual a la concentración del contaminante medida.
- BP_{L_0} := Es el punto de corte menor o igual a la concentración del contaminante medida.
- I_{Hi} := Es el valor del *ICA* correspondiente al punto de corte de BP_{Hi} .
- I_{L_0} := Es el valor del *ICA* correspondiente al punto de corte BP_{L_0} .

Marco Normativo y Protocolo

La resolución 2254 del ministerio de ambiente y desarrollo sostenible del 01 de noviembre del 2017 por la cual se adopta la norma de calidad del aire ambiente y se dictan otras disposiciones estable en el artículo 2 los niveles máximos permisibles de contaminantes críticos

Artículo 2 En la tabla (3-1) se establecen los niveles máximos permisibles a condiciones de referencia para contaminantes. Criterios que regirán a partir de enero del año 2018

Tabla 3-1: Niveles a partir de 2018

Contaminante	Nivel máximo Permissible ($\mu\text{g}/\text{m}^3$)	Tiempo de exposición
PM ₁₀	50	Anual
	100	24 horas
PM _{2.5}	25	Anual
	50	24 horas
SO ₂	50	24 horas
	100	1 hora
NO ₂	60	Anual
	200	1 hora
O ₃	100	8 horas
CO	5.000	8 horas
	35.000	1 hora

Tomado de <http://www.minambiente.gov.co/images/normativa/app/resoluciones/96-res%202254%20de%202017.pdf>

De igual manera en el artículo 3 se establecen los niveles máximos permisibles de contaminantes en el aire para el año 2030, los niveles de contaminación se observan en la imagen (3-2).

Tabla 3-2: Niveles máximos permisibles de contaminantes en el aire para el año 2030

Contaminante	Nivel máximo Permissible ($\mu\text{g}/\text{m}^3$)	Tiempo de exposición
PM ₁₀	30	Anual
PM _{2.5}	15	Anual
SO ₂	20	24 horas
NO ₂	40	Anual

Según el Sistema de Información Ambiental de Colombia (SIAC) las principales emisiones de contaminantes al aire del país son: el Material Particulado - MP, los óxidos de

azufre SO_2 , los óxidos de nitrógeno NO_2 , el monóxido de carbono CO (contaminantes criterio) y el dióxido de carbono CO_2 .

3.1.1. Material Particulado (PM)

El material particulado es un indicador de contaminación del aire que afecta a la gran parte de la población (humana, animales y plantas). Los componentes principales de PM son los sulfatos, los nitratos, el amoníaco, el cloruro de sodio, el hollín, los polvos minerales y el agua. Consisten en una compleja mezcla de partículas sólidas y líquidas de sustancias orgánicas e inorgánicas suspendidas en el aire [O.M.S. \(2018\)](#)

- **PM₁₀** Consiste en una partícula con diámetro aerodinámico equivalente igual a menos a 10 micrómetros (Bracho, 2013). Esta partícula es sólida o líquida de polvo y está presente en fuentes móviles y fijas de procesos mecánicos que se puede encontrar en las cenizas, hollín, partículas metálicas, cemento o polen, dispersa en la atmósfera. La exposición prolongada o repetitiva a la PM10 puede producir efectos nocivos en el sistema respiratorio de la persona [Galvis y Rojas \(2006\)](#).

- **PM_{2.5}**

El PM_{2.5} es una partícula con diámetro aerodinámico inferior o igual a los 2,5 micrómetros. Esta partícula al igual que PM₁₀ corresponde a fuentes finas de combustiones, incendios forestales, y algunos procesos industriales. La exposición prolongada a la PM_{2.5} puede producir enfermedades cardiovasculares, respiratorias y cáncer (Jiménez, 2008).

3.1.2. Gases contaminantes

Los gases contaminantes presentes en la atmósfera corresponde al ozono (O_3), dióxido de nitrógeno (NO_2), dióxido de azufre (SO_2) y monóxido de carbono (CO). Estos gases contaminantes son producidos por fuentes móviles y fuentes fijas de combustión. Se clasifican en gases contaminantes primarios los cuales producen a partir de fuentes de emisión, y gases contaminantes secundarios con aquellos que se producen por la reacción física y química que se generan a partir de los contaminantes primarios

- **Ozono (O_3)** El O_3 es una de las varias configuraciones químicas en las que el oxígeno se encuentra presente en la naturaleza. El O_3 se caracteriza por ser un gas de color azul pálido que se encuentra en la parte superior de la atmósfera de la estratósfera, de gran ayuda para los seres vivos (Humanos, animales y plantas) siendo una capa protectora de los rayos ultravioleta. También encontramos el ozono troposférico presente en partes bajas de la atmósfera, el cual es perjudicial para los seres vivos, si la encontramos en grandes concentraciones. La presencia de este tipo de O_3 puede

afectar el sistema respiratorio el cual ataca al tejido pulmonar, también genera afectación en las plantas (Geoambiental, 2018).

- Dióxido De Nitrógeno (NO_2)

El NO_2 corresponde a un compuesto químico gaseoso de color pardo rojizo, correspondiente a la combinación de un átomo de nitrógeno y dos de oxígeno. El NO_2 hace presencia en el aire específicamente en aquellas zonas urbanas con mayor concentración de vehículos (combustión en diésel) e industria. El NO_2 es también producido por incendios forestales o erupciones volcánicas, de forma natural en la descomposición de nitratos orgánicos. Es una fuente de material particulado $\text{PM}_{2.5}$ especialmente, y al contacto prolongado a la luz UV del sol es un buen iniciador de O_3 ozono troposférico. La exposición prolongada a este gas afecta a los seres vivos en los pulmones y en las vías respiratorias (SDA, 2017).

- Dióxido de Azufre SO_2

Corresponde a un Contaminante primario que se caracteriza por ser un gas incoloro y con un olor picante e irritante. También es conocido como dióxido de sulfuro, no es un gas explosivo ni inflamable, es salubre en el agua y al combinarse se convierte en ácido sulfúrico. El dióxido de azufre se produce a partir de combustibles fósiles que contienen carbón, gasolina y diésel, presentes en fuentes móviles y fijas (Centrales eléctricas). Al contacto con otros componentes químicos se puede producir lluvia ácida y material particulado $\text{PM}_{2.5}$. La exposición prolongada a SO_2 puede generar graves problemas para la salud especialmente al sistema circulatorio, inflamación en las vías respiratorias y paro cardíaco

- Monóxido De Carbono (CO).

El monóxido de carbono es un contaminante primario que se caracteriza por ser inodoro, incoloro e inflamable, es parcialmente soluble en el agua. El CO es producido por la combustión incompleta de fósiles como gas natural, gas propano, gasolina, madera, carbón y petróleo. La exposición prolongada al CO produce daños en el transporte de oxígeno en la sangre produciendo la muerte

Variables meteorológicas

En el estudio de la atmósfera se clasifica en meteorología y la climatología. La meteorología se encarga en explicar, analizar y predecir los fenómenos físicos que desarrolla la atmósfera. La climatología se encarga de la variable meteorológica la cual trata de documentar, analizar y explicar las variables espaciales implicando diferentes escalas temporales (mensuales, anuales, estacionales... etc) (Nadal (2006))

Capítulo 4

Marco Metodológico

4.1. Fuentes de información

Las fuentes de información principalmente fueron las siguientes:

- Los datos utilizados en el desarrollo de este trabajo tienen como fuente los datos publicados de las estaciones de la Red de Monitoreo de Calidad del Aire de Bogotá - RMCAB, que se ingresa obtienen a través de la página web de la secretaria distrital de ambiente .
- Los datos utilizados corresponden al intervalo de tiempo entre el 04 de Abril de 2019 y el 22 de Abril de 2019 para cuatro estaciones de Monitoreo: Carvajal-Sevilla, Centro de Alto rendimiento, Tunal y Usaquen de los atmosféricos: SO₂, NO₂, CO, O₃.

La ciudad de Bogotá cuenta con la Red de Monitoreo de Calidad del Aire de Bogotá denominada RMCAB. El año 2016 RMCAB contaba con trece estaciones las cuales doce de ella recolectaban información sobre la concentración de contaminantes y variables meteorológicas, y la estación Bolivia encargada de hacer recolección de información únicamente de las variables meteorológicas. La información de la Red de Monitoreo de Calidad del Aire de Bogotá denominada RMCAB mensual, las estaciones se encuentran ubicadas de manera estratégica con el fin de obtenga la mayor cobertura de la ciudad, esta información es procesada en la estación principal en donde se valida y se evalúa a partir de los estándares de calidad del aire en Bogotá (Ambiente, 2018).

Las estaciones recogen información de la siguiente manera:

- Material particulado (PM₁₀, PM_{2.5})
- Gases contaminantes (SO₂, NO₂, CO, O₃)

De igual manera recogen información de las variables meteorológicas:

- Precipitación
- Velocidad y Dirección del Viento
- Temperatura
- Radiación Solar
- Humedad Relativa
- Presión Barométrica que regulan el transporte de los contaminantes en la atmósfera de Bogotá.

Tabla 4-1: Ubicación estaciones RM CAB

Estación	Latitud	Longitud	Altitud	Localidad	dirección	Tipo de zona	Tipo de estación
Usaquén	442°37.26' N	741°49.50' W	2570 m	Usaquén	Carrera 7B Bis # 132-11	Urbana	De fondo
Tunal	434°34.41' N	747°51.44' W	2589 m	Tunjuelito	Carrera 24 # 49-86 sur	Urbana	De fondo
Carvajal	435°44.22' N	748°54.90' W	2563 m	Kennedy	Autopista Sur # 63-40	Urbana	Tráfico Industrial
P. Simón Bolívar	439°30.48' N	745°2.28' W	2577 m	Barrios Unidos	Calle 63 # 59A-06	Urbana	De fondo
Puente Aranda	437°54.36' N	747°2.94' W	2590 m	Puente Aranda	Calle 10 # 65-28	Urbana	Industria
Kennedy	437°30.18' N	749°40.80' W	2580 m	Kennedy	Carrera 80 # 40-55 sur	Urbana	De fondo

En Colombia la entidad gubernamental que vigila de la calidad del aire se llama El IDEAM (Instituto de Hidrología, Meteorología y Estudios Ambientales), es una institución de carácter pública que apoya técnico y científico al Sistema Nacional Ambiental, que genera conocimiento, produce información confiable, consistente y oportuna, sobre el estado y las dinámicas de los recursos naturales y del medio ambiente, que facilite la definición y ajustes de las políticas ambientales y la toma de decisiones por parte de los sectores público, privado y la ciudadanía en general. (IDEAM, 2018)

Capítulo 5

Análisis y Resultados

5.1. Técnica de análisis de datos

5.1.1. Datos de Panel

Segun [Wooldridge y cols. \(2009\)](#), un conjunto de datos de panel (o longitudinales) consiste en una serie de tiempo por cada unidad de una base de datos de corte transversal.

Los datos de panel, son definidos también [Perez \(s.f.\)](#) como un conjunto de variables para un conjunto de unidades sociales o secciones, cuyos valores se observan durante un periodo de tiempo. Un conjunto de datos panel presenta tres dimensiones: unidad de análisis o agente social, variables y tiempo, la particularidad de esta estructura se presenta en que la muestra tomada permanece constante a lo largo de los periodos de tiempo observados.

El objetivo de estudios que emplean información tipo panel de datos es identificar heterogeneidad no observable, analizando los efectos individuales específicos y los efectos temporales ya que cuenta con una dimensión temporal de los datos.

5.1.2. Especificación general de un modelo de datos de panel

De acuerdo a [Baronio y Vianco \(2014\)](#) la especificación general de un modelo de regresión con datos de panel es la siguiente:

$$Y_{it} = a_{it} + b_1 X_{1it} + b_2 X_{2it} + \dots + b_k X_{kit} + U_{it}; \quad i = 1, \dots, n \quad y \quad t = 1, \dots, T \quad (5-0)$$

Donde:

- i se refiere al individuo o a la unidad de estudio (corte transversal)
- t a la dimensión en el tiempo.
- a es un vector de interceptos que puede contener entre 1 y $n + 1$ parámetros.

- b es un vector de k parámetros.
- X_{it} es la i -ésima observación al momento t para las variables explicativas X_1, X_2, \dots, X_k

La muestra total de las observaciones en el modelo viene es $n \times T$

A continuación se muestra la forma de la tabla de datos

Tabla 5-1: Tabla Datos panel

Tiempo	Individuo	Y	X_1	X_2	\dots	X_k
1	A		X_{1A1}	X_{2A1}		
2	A		X_{1A2}	X_{2A2}		
\vdots			\vdots	\vdots	\vdots	\vdots
T	A		X_{1AT}	X_{2AT}		
1	B		X_{1B1}	X_{2B1}		
2	B		X_{1B2}	X_{2B2}		
\vdots	\vdots		\vdots	\vdots	\vdots	\vdots
T	B		X_{1BT}	X_{2BT}		
1	C		X_{1C1}	X_{2C1}		
2	C		X_{1C2}	X_{2C2}		
\vdots	\vdots		\vdots	\vdots	\vdots	\vdots
T	C		X_{1CT}	X_{2CT}		

5.1.3. Componentes del error

Es usual interpretar los modelos de datos de panel a través de sus componentes de errores. El término de error U_{it} o en la ecuación (5.1.2) se descompone como:

$$U_{it} = u_i + v_t + w_{it} \quad (5-0)$$

- u_i representa los efectos no observables que difieren entre las unidades de estudio pero no en el tiempo, que generalmente se los asocia a la tecnología incorporada.
- v_t son los efectos no cuantificables que varían en el tiempo pero no entre las unidades de estudio.
- w_{it} se refiere al término de error puramente aleatorio.

5.1.4. Ventajas y Desventajas de la Técnica de Datos de Panel

La técnica de datos tipo panel presenta una serie de ventajas y desventajas en comparación con los modelos de series de tiempo y de corte transversal.

Ventajas:

- La técnica permite al investigador disponer de un mayor número de observaciones $n \times T$ incrementando los grados de libertad, reduciendo la colinealidad entre las variables explicativas y, en última instancia, mejorando la eficiencia de las estimaciones.
- La técnica permite capturar la heterogeneidad no observable ya sea entre unidades individuales de estudio como en el tiempo. Con base en lo anterior, la técnica permite aplicar una serie de pruebas de hipótesis para confirmar o rechazar dicha heterogeneidad y cómo capturarla.
- Los datos en panel suponen e incorporan, en el análisis los individuos. Los análisis de series de tiempo y de corte transversal no tratan de controlar esta heterogeneidad corriendo el riesgo de obtener resultados sesgados.
- Permite elaborar y probar modelos relativamente complejos de comportamiento en comparación con los análisis de series de tiempo y de corte transversal.

Desventajas: En términos generales, las desventajas asociadas a la técnica de datos de panel se relacionan con los procesos para la obtención y el procesamiento de la información estadística sobre las unidades individuales de estudio, cuando esta se obtiene por medio de encuestas, entrevistas o utilizando algún otro medio de relevamiento de los datos.

Modelo de Efectos fijos

El modelo de efectos fijos considera que existe un término constante diferente para cada individuo y supone que los efectos individuales son independientes entre sí. Con este modelo se considera que las variables explicativas afectan por igual a las unidades de corte transversal y que éstas se diferencian por características propias de cada una de ellas, medidas por medio del intercepto. Es por ello que los interceptos se asocian con variables dummy con coeficientes específicos para cada unidad, los cuales se deben estimar. Para la i -ésima unidad de corte transversal, la relación es la siguiente:

$$Y_i = ia_i + b_1X_{1i} + b_2X_{2i} + \dots + b_kX_{ki} + U_i; \quad i = 1, 2, \dots, n \quad (5-0)$$

Donde el subíndice i representa un vector columna de unos individuos. Este modelo presenta los residuos

$$U_{it} = u_i + v_t + W_{it}$$

Siendo w_{it} , variable aleatoria y u_i (efectos de sección cruzada) y v_t (efectos de tiempo) deterministas.

Modelo de Efectos Aleatorios

Según [Mayorga y Evelyn Muñoz \(2000\)](#) el modelo de efectos aleatorios considera que los efectos individuales no son independientes entre sí, sino que están distribuidos aleatoriamente alrededor de un valor dado. El modelo considera que tanto el impacto de las variables explicativas como las características propias de cada unidad de corte transversal son diferentes. El modelo se expresa algebraicamente de la siguiente forma:

$$Y_{it} = (a + u_i) + b_1X_{1it} + b_2X_{2it} + \dots + b_kX_{kit} + w_{it}; \quad i = 1, 2, \dots, n \quad t = 1, 2, \dots, T \quad (5-0)$$

Test de Hausman El test propuesto por Hausman en el año 1978, es un test Chi cuadrado que determina si las diferencias son sistemáticas y significativas entre dos estimaciones, comparando los β obtenidos por medio del estimador de efectos fijos y efectos aleatorios, identificando si las diferencias entre ellos son o no son significativas. En la hipótesis nula se comprueba la existencia de no correlación entre los α_i y las variables explicativas.

H_0 : = *No existe correlación entre las variables explicativas y los efectos individuales.*
 H_1 := *Existe correlación entre las variables explicativas y los efectos individuales.*

Criterio de rechazo

- Si el p -valor < 0.05 se rechaza H_0 , es decir que hay correlación entre los efectos individuales y las variables explicativas, lo que indica que el estimador fijo debe ser utilizado.
- Si el p -valor > 0.05 , no se rechaza H_0 , es decir que no hay una correlación entre los efectos individuales y las variables explicativas, lo que indica que el estimador de efectos aleatorios es el que se debe utilizar.

5.2. Resultados

En las sección anterior se fundamentó la parte teórica que sustentan la estimación de un modelo datos panel, por lo tanto en esta sección se mostraran los resultados de la aplicación de los modelos iniciando con una contextualización de la información a hasta los resultados de la estimación

En esta sección se muestran los resultados del análisis hecho sobre los gases contaminantes Dióxido de Azufre (SO_2), y Ozono O_3 .

5.2.1. Panel de datos Dióxido de Azufre

Como ya se menciona en el marco metodológico Los datos utilizados para la aplicación de este trabajo tienen como fuente los datos publicados de las estaciones de la Red de Monitoreo de la Calidad del Aire de Bogotá (RMCAB) obtenidos a través de la página web de la secretaria distrital de ambiente Para la aplicación del modelo se tomó la base de datos discriminada por las principales para 4 zonas de Bogotá y por un total de 19 periodos diarios comenzando en 04-04-2019 y finalizando el 22-04-2019, las variables a tener en cuenta serán

Gases contaminantes (SO_2 , NO_2 , CO , O_3), Material particulado (PM_{10} , $\text{PM}_{2.5}$) y De igual manera recogen información de las variables meteorológicas: Precipitación Velocidad y Dirección del Viento, Temperatura

- **Variable Respuesta (Dióxido de azufre SO_2):** Cantidad de Niveles de concentración medida en partículas por billon (*ppb*) del contaminante atmosférico Dióxido de Azufre SO_2 .
- Las 4 zonas de Bogotá a tener en cuenta son: Carvajal-Savillana, Tunal, Centro de Alto Rendimiento y Usaquén.

De esta manera, en la tabla (5-2) se observan algunas estadísticas descriptivas para la variable de respuesta SO_2 como lo son la media, la media, la mediana y la desviación estándar.

Tabla 5-2: Estadísticas Descriptivas

Descriptive Statistics for SO_2
Categorized by values of SO_2
Sample: 4/04/2019 4/22/2019
Included observations: 76

SO_2	Mean	Median	Std. Dev.	Obs.
[0, 5)	1.223793	1.150000	0.678972	58
[5, 10)	7.511111	7.900000	1.398610	9
[10, 15)	12.54286	12.50000	1.126731	7
[15, 20)	17.35000	17.35000	0.353553	2
All	3.435263	1.400000	4.398050	76

En la figura (5-1) se muestran las secciones transversales combinadas la cual muestra líneas separadas para cada sección transversal, en este caso cada una de las 4 zonas, se puede visualizar que para zona Carvajal-Sevillana los niveles de dióxido de Azufre SO_2 son más altos que para cada una de las otras tres zonas.

Según el artículo " *Así se vive en La Sevillana, el segundo lugar más contaminado de Colombia*" Un ejemplo, ese fue el único lugar, de todos los medidos en el país, donde la presencia de dióxido de azufre en el ambiente no estuvo dentro de la categoría "buena". Esto porque en este sector confluyen grandes fábricas, un río podrido, el matadero más grande de la zona y dos de las vías de mayor congestión, recorridas por tractomulas y camiones a toda hora.

Figura 5-1: Secciones transversales combinadas

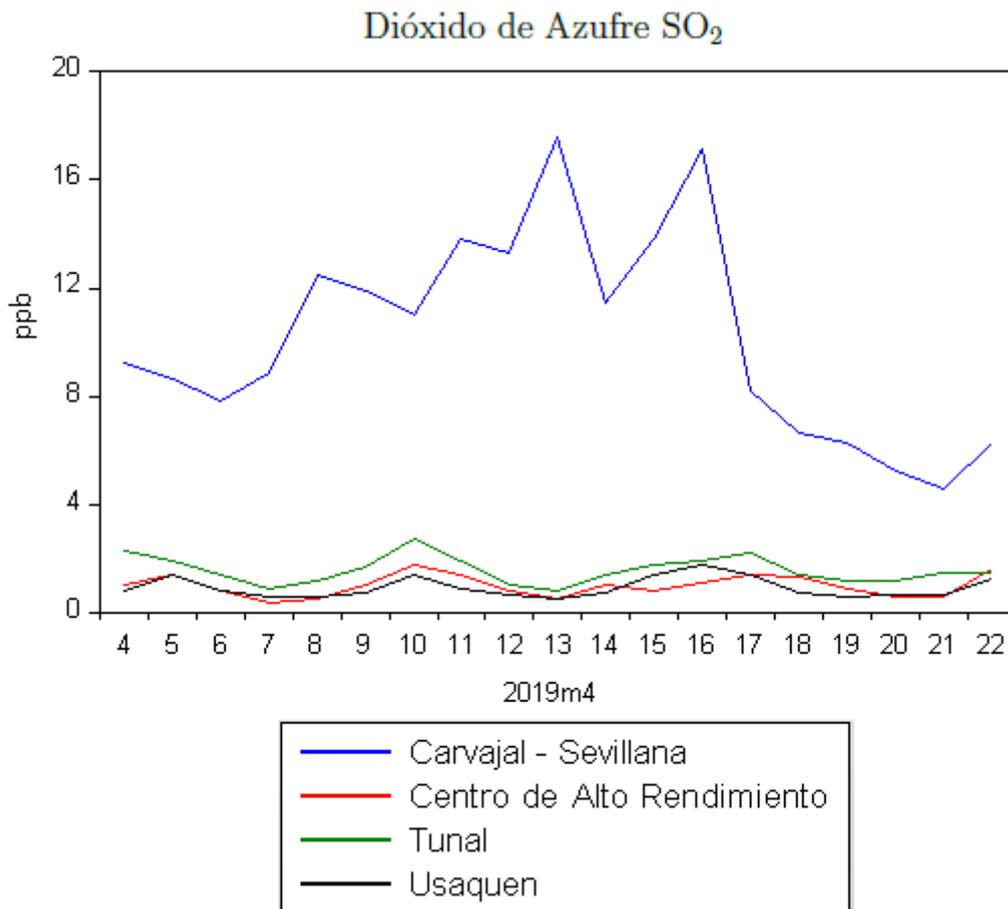
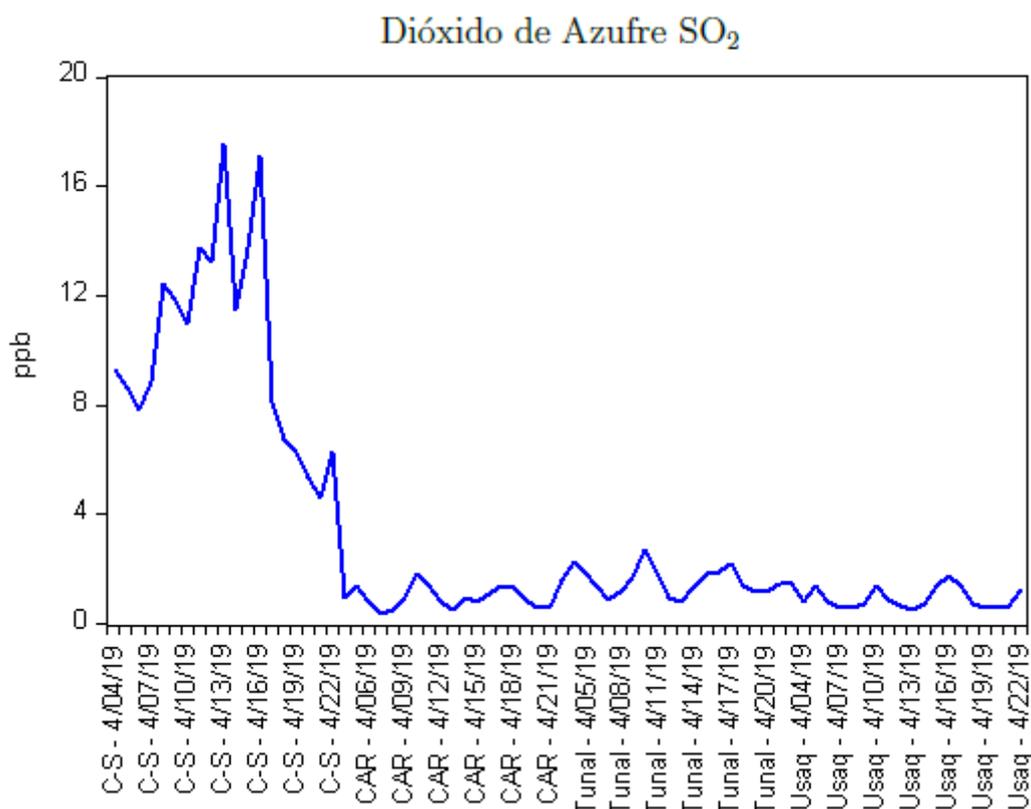


Figura 5-2: Secciones Transversales apiladas



En la figura (5-2) se muestra un gráfico de los datos apilados, etiquetados con la sección transversal y la fecha.

Para continuar con la podido seguir con el estudio de las variables por medio del modelo de datos de panel de efectos fijos es necesario realizar la prueba de medias.

a) Prueba de igualdad de medias.

Tabla 5-3: Prueba igualdad de medias.

Test for Equality of Means of SO ₂			
Method	df	Value	Probability
Anova F-test	(3, 72)	676.9943	0.0000
Welch F-test*	(3, 4.65847)	1100.343	0.0000

*Test allows for unequal cell variances

Para el contraste de igualdad de medias de la variable SO₂, se realiza la prueba de

igualdad de medias cuyos resultados se observan la tabla **5-3** y lo que muestra que se rechaza la igualdad de medias (p - valores menores que 0.05) Lo mismo se repite hace para todas las variables.

b) Prueba de igualdad de varianzas.

Tabla **5-4**: Prueba igualdad de varianzas.

Test for Equality of Variances of SO ₂			
Method	df	Value	Probability
Bartlett	3	11.08835	0.0113
Levene	(3, 72)	7.394109	0.0002
Brown-Forsythe	(3, 72)	5.791900	0.0013

Para el contraste de igualdad de varianzas de la variable SO₂, se realiza la prueba de igualdad de varianzas cuyos resultados se pueden observar la tabla **5-4** donde se concluye para cada una de las pruebas que se rechaza la igualdad de varianzas pues los (p - valores son menores que 0.05). Esta prueba se realiza para cada una de las variables.

c) Estimación del modelo. **Efectos fijos**

Tabla **5-5**: Estimación del modelo

Dependent Variable: SO ₂				
Method: Panel Least Squares				
Sample: 4/04/2019 4/22/2019				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DIR_VIENTO	-0.005224	0.002216	-2.357868	0.0223
PM2_5	0.334493	0.039377	8.494598	0.0000
VEL_VIENTO	0.897564	0.525468	1.708123	0.0937
C	-5.281104	1.325445	-3.984401	0.0002

En la tabla (**5-5**) se muestran las variables que resultaron significativas para explicar la variable de respuesta (dióxido de Azufre), estas fueron la dirección y velocidad del viento así como el material particulado PM_{2.5} tomando como referencia p -valores menores a 0.1 En este modelo se puede observar que cuando aumenta tanto la velocidad del viento como el material particulado PM_{2.5} aumenta los niveles de concentración del

dióxido de azufre, mientras que si hay un aumento (cambio en la dirección del viento) hay una pequeña disminución en los niveles de dióxido de azufre.

Tabla 5-6: R^2 y Durbin -Watson

Effects Specification			
Cross-section fixed (dummy variables)			
Period fixed (dummy variables)			
R-squared	0.950386	Mean dependent var	3.435263
Adjusted R-squared	0.927038	S.D. dependent var	4.398050
S.E. of regression	1.187980	Akaike info criterion	3.441373
Sum squared resid	71.97612	Schwarz criterion	4.208061
Log likelihood	-105.7722	Hannan-Quinn criter.	3.747779
F-statistic	40.70540	Durbin-Watson stat	1.659381
Prob(F-statistic)	0.000000		

Para el análisis de los resultados se toma el valor del coeficiente de correlación R^2 y el valor de Durbin – Watson, los cuales obtuvieron un valor 0.950386 y 1.659381 respectivamente, lo que indica que el modelo presenta buena calidad para replicar los resultados, y expresa el 95,03 % de la varianza de la variable dependiente (SO_2) explicada por las variables independientes que en este caso con la velocidad del viento y la variable de material particulado $PM_{2.5}$ y el valor de Durbin – Watson nos dice que los residuos no están correlacionados.

d) Validación del modelo

A continuación se muestran los resultados que nos permiten validar el modelo.

- I. De acuerdo con la tabla (5-6) $R^2 = 0.950386$
- II. De la tabla (5-6) Durbin-Watson stat:= 1.659381
- III. Normalidad.

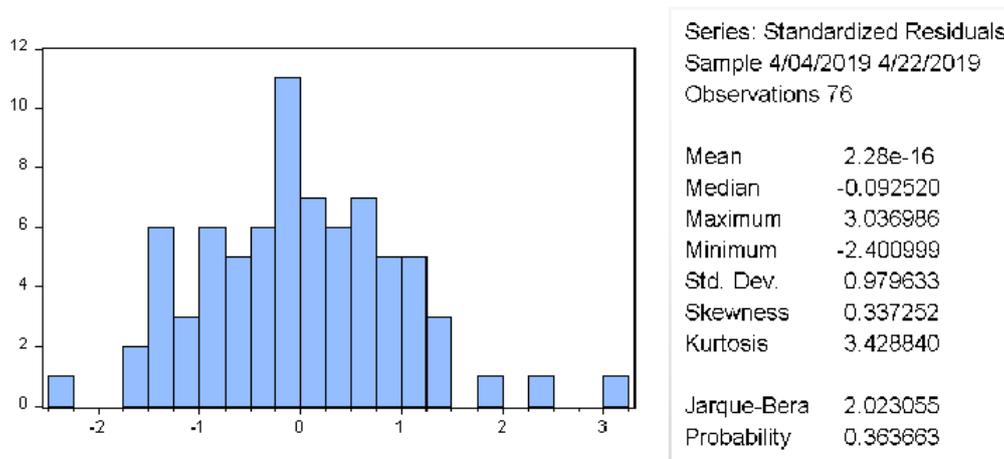
Contraste de hipótesis

H_0 : Los residuales son normales.

H_1 : Los residuales no son normales.

Se observa en la figura (5-3) que el p -valor del contraste de normalidad de Jarque-Bera es superior a 0.05; lo que indica residuos normales con una confianza del 95 %.

Figura 5-3: Prueba de Normalidad



IV. Para probar si los efectos fijos de las estaciones y de tiempo pueden o no considerarse iguales se utiliza el test de máxima verosimilitud para la redundancia. En la tabla (5-7) se observan p -valores menores que 0.05, lo que nos lleva a afirmar que los efectos fijos de las estaciones y de tiempo son diferentes con un 95% de confianza.

Tabla 5-7: Add your caption here.

Redundant Fixed Effects Tests

Test cross-section and period fixed effects

Effects Test	Statistic	d.f.	Prob.
Cross-section F	8.725407	(3,51)	0.0001
Cross-section Chi-square	31.484196	3	0.0000
Period F	3.438329	(18,51)	0.0003
Period Chi-square	60.388658	18	0.0000
Cross-Section/Period F	16.331891	(21,51)	0.0000
Cross-Section/Period Chi-square	155.378080	21	0.0000

V. Prueba de dependencia de los residuales

Las pruebas de dependencia de sección transversal se muestran en la tabla (5-8)

Tabla 5-8: Prueba dependencia de los residuales

Residual Cross-Section Dependence Test
 Null hypothesis: No cross-section dependence (correlation) in residuals

Test	Statistic	d.f.	Prob.
Breusch-Pagan LM	21.12965	6	0.0017
Pesaran scaled LM	4.367555		0.0000
Bias-corrected scaled LM	4.256444		0.0000
Pesaran CD	-3.034848		0.0024

La primera línea contiene los resultados de la prueba Breusch-Pagan LM. En este caso, el valor del estadístico de prueba, 21.12965, y rechazamos hipótesis nula de no correlación en los niveles de significación convencionales.

Las siguientes líneas presentan resultados para las dos pruebas de Breusch-Pagan escaladas y para Pesaran las cuales rechazan fuertemente la hipótesis nula a un 5%.

VI. Heterocedasticidad

Tabla 5-9: Heterocedasticidad de los residuales

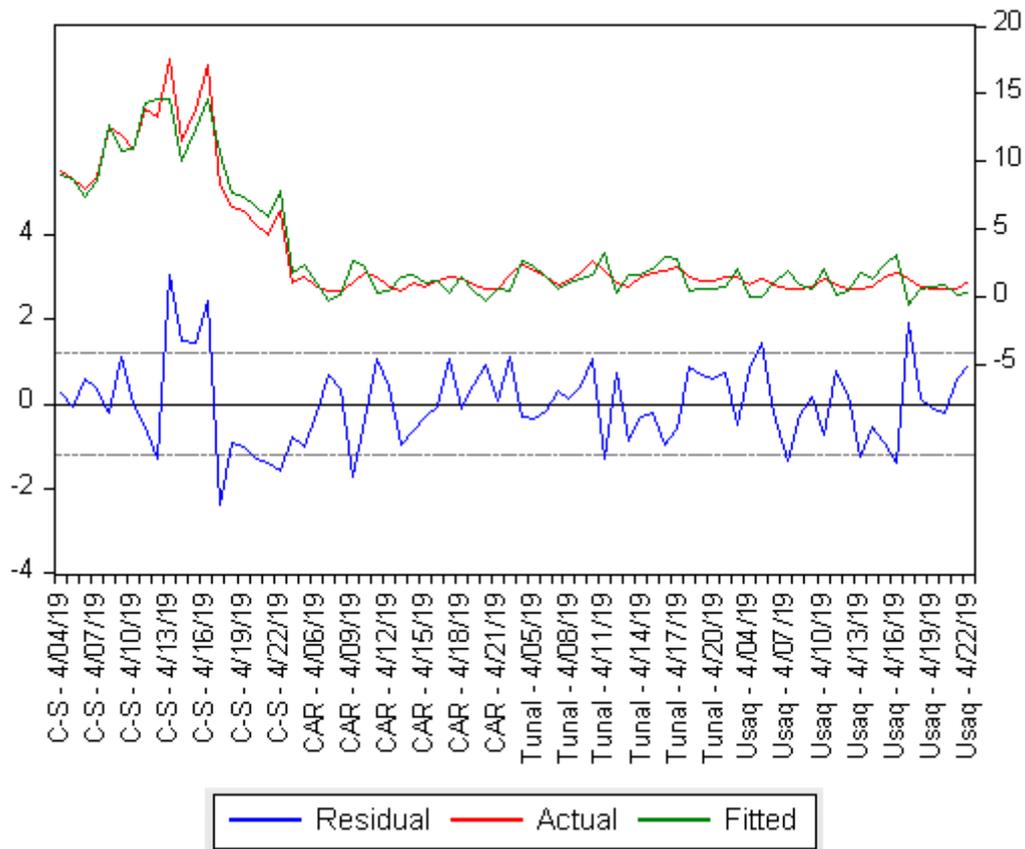
Test for Equality of Variances of RESID
 Categorized by values of RESID

Method	df	Value	Probability
Bartlett	3	0.104587	0.9913
Levene	(3, 72)	1.271776	0.2906
Brown-Forsythe	(3, 72)	1.038639	0.3808

La tabla (5-9) muestra los resultados, observándose que se acepta la igualdad de varianzas residual en las distintas secciones cruzadas dado que los p -valores en cada una de las pruebas con mayores que 0.05. No existe entonces heteroscedasticidad.

- e) Gráfico de estimación En la figura (5-4) se muestra el gráfico de estimación del dióxido de azufre.

Figura 5-4: Gráfico de estimación



f) Efectos individuales - Efectos tiempo

A continuación se muestran los efectos de sección transversal por estaciones.

Tabla 5-10: Cross-section Fixed effects SO₂

INDIVIDUO	Effect
C-S	2.710304
CAR	-0.996468
Tunal	-0.888751
Usaq	-0.825085

En la tabla (5-10) se evidencia que en los sectores de Usaquén, Centro de alto rendimiento y Tunal hay efecto negativo, mientras que en el sector de Carvajal Sevillana hay un efecto positivo.

Tabla 5-11: Period Fixed Effects

FECHA	Effect
2019 - 04 - 04	1.285758
2019 - 04 - 05	-1.972454
2019 - 04 - 06	0.253160
2019 - 04 - 07	-2.259868
2019 - 04 - 08	-4.075923
2019 - 04 - 09	-2.226071
2019 - 04 - 10	-1.579035
2019 - 04 - 11	-2.743726
2019 - 04 - 12	-1.262568
2019 - 04 - 13	-0.810302
2019 - 04 - 14	0.111208
2019 - 04 - 15	2.087144
2019 - 04 - 16	1.730995
2019 - 04 - 17	-0.370515
2019 - 04 - 18	1.882449
2019 - 04 - 19	3.479396
2019 - 04 - 20	2.822688
2019 - 04 - 21	3.042462
2019 - 04 - 22	0.605203

En la tabla (5-11) se observa que los días de mayor y menor aporte fueron el 2019-04-19 y el 2019-04-08 respectivamente

g) Ecuación para estimar los niveles de SO_2

Según los resultados obtenidos en los anteriores items a continuación se presenta la ecuación del modelo de efectos fijos ajustado para la estimación de la variable de respuesta, es decir para el dióxido de Azufre

$$SO_2 = -5.81104 - 0.005224 \cdot dir\ viento + 0.334493PM_{2.5} + 0.897564velviento + 2.710304d_1 - 0.996468d_2 - 0.888751d_3 - 0.825085d_4 + 1.285758A_4 - 1.972454A_5 + 0.253160A_6 - 2.259868A_7 - 4.075923A_8 - 2.226071A_9 - 1.579035A_{10} - 2.743726A_{11} - 1.262568A_{12} - 0.810302A_{13} + 0.111208A_{14} + 2.087144A_{15} + 1.730995A_{16} - 0.370515A_{17} + 1.882449A_{18} + 3.479396A_{19} + 2.822688A_{20} + 3.042462A_{21} + 0.605203A_{22}$$

($d_i = 1$ para observaciones de la estación i y vale $d_i = 0$ en caso contrario, de la misma forma $A_t = 1$ para observaciones del día t y $A_t = 0$ en caso contrario)

5.2.2. Panel de datos Ozono

En esta sección se considera la variable de respuesta a la variable ozono (O_3). que indica los niveles de concentración medida en *ppb*

De manera análoga a lo que se hizo con el dióxido de azufre se empiezan mostrando algunas estadísticas descriptivas en la tabla (5-12) como lo son la media la mediana y la desviación estándar.

Tabla 5-12: Estadísticas descriptivas ozono

Descriptive Statistics for OZONO Categorized by values of OZONO				
OZONO	Mean	Median	Std. Dev.	Obs.
[0, 10)	8.344000	8.800000	1.338868	25
[10, 20)	14.31935	14.10000	2.641643	31
[20, 30)	23.77333	23.60000	2.149574	15
[30, 40)	35.46000	34.10000	3.104513	5
All	15.61053	13.60000	7.913559	76

En la gráfica (5-5) se muestran las secciones transversales combinadas la cual muestra líneas separadas para cada sección transversal que corresponde a cada una de las 4 estaciones de monitoreo de la calidad del aire en Bogotá que se han seleccionado para el estudio, se observa que para zona de Usaquén los niveles de Ozono O_3 son más altos que para cada una de las otras tres zonas.

En la figura (5-6) se muestra un gráfico de los datos apilados, etiquetados con la sección transversal y la fecha. Con un gráfico de tipo de línea y símbolo , en este gráfico se pudo observar que los niveles de ozono fueron más altos en Usaquén, y más bajos en La sevillana, mientras que en el centro de alto rendimiento y Tunal fueron muy similares.

Figura 5-5: Secciones transversales O_3

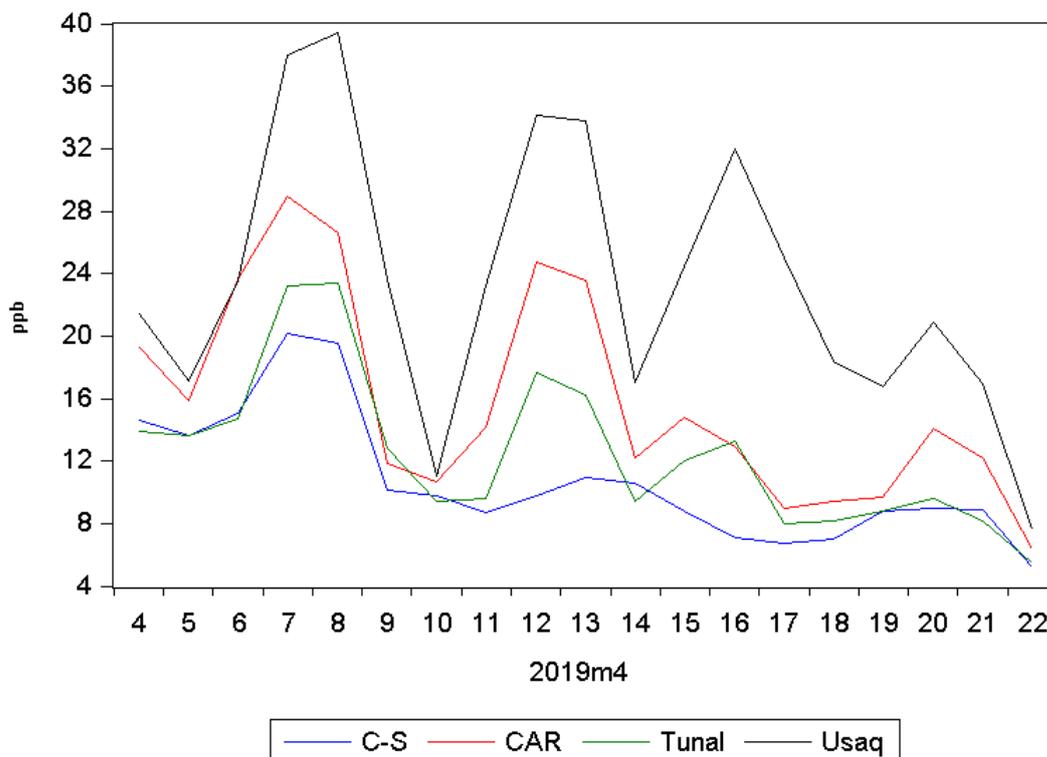
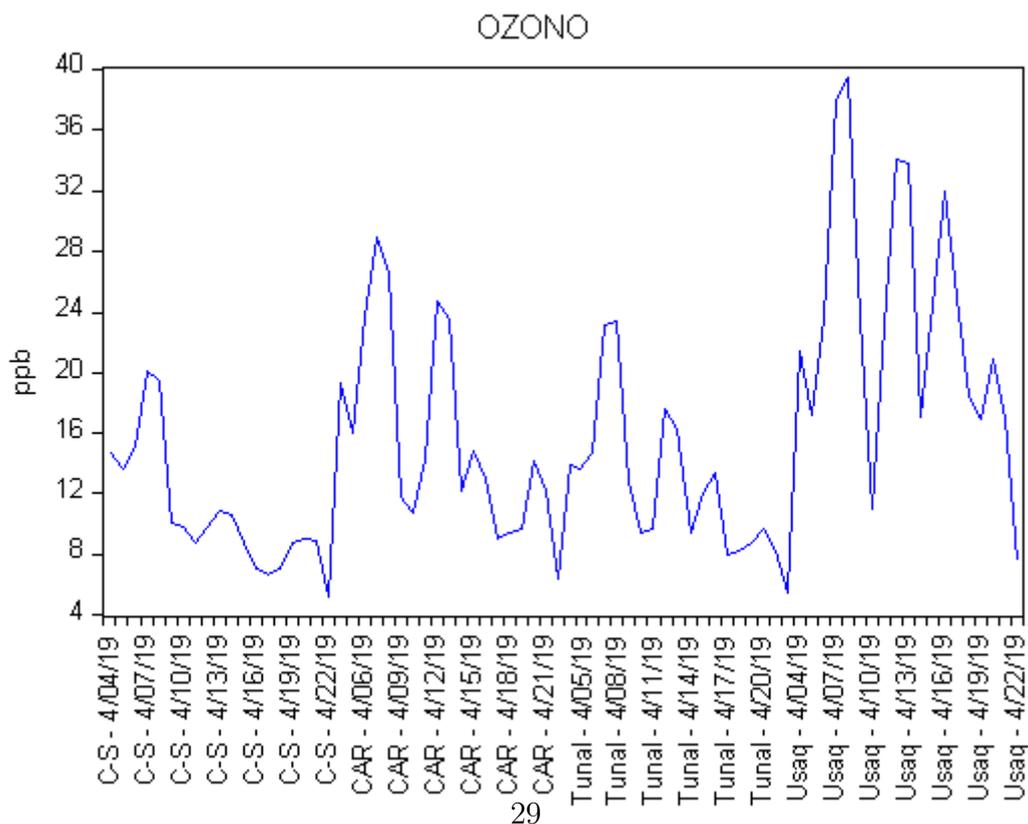


Figura 5-6: Secciones transversales apiladas O_3



a) Prueba de igualdad de medias para el ozono

Tabla 5-13: Igualdad de medias.

Test for Equality of Means of OZONO			
Method	df	Value	Probability
Anova F-test	(3, 72)	292.9864	0.0000
Welch F-test*	(3, 15.7489)	289.3537	0.0000

*Test allows for unequal cell variances

b) Prueba igualdad de varianzas.

De acuerdo a la tabla (5-14)

Tabla 5-14: Igualdad de varianzas Ozono

Test for Equality of Variances of OZONO			
Method	df	Value	Probability
Bartlett	3	11.71637	0.0084
Levene	(3, 72)	4.356488	0.0071
Brown-Forsythe	(3, 72)	3.482104	0.0201

c) Estimación del modelo. **Modelo de efectos fijos**

En la tabla (5-15) se muestran las variables que resultaron significativas para explicar la variable de respuesta (Ozono), estas fueron la velocidad del viento así como el material particulado $PM_{2.5}$ tomando como referencia p -valores menores a 0.056 En este modelo se puede observar que cuando aumenta la velocidad del viento aumenta los niveles de concentración del Ozono, mientras que si hay un aumento en los niveles de material particulado $PM_{2.5}$ hay una disminución en los niveles de Ozono.

Tabla 5-15: Add your caption here.

Dependent Variable: OZONO
 Method: Panel Least Squares
 Cross-sections included: 4
 Total panel (balanced) observations: 76

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PM2_5	-0.418676	0.095147	-4.400310	0.0001
VEL_VIENTO	2.954323	1.174546	2.515291	0.0150
C	22.51066	2.781105	8.094143	0.0000

Effects Specification

Cross-section fixed (dummy variables)
 Period fixed (dummy variables)

R-squared	0.908663	Mean dependent var	15.61053
Adjusted R-squared	0.868264	S.D. dependent var	7.913559
S.E. of regression	2.872262	Akaike info criterion	5.200166
Sum squared resid	428.9943	Schwarz criterion	5.936187
Log likelihood	-173.6063	Hannan-Quinn criter.	5.494316
F-statistic	22.49220	Durbin-Watson stat	1.458439
Prob(F-statistic)	0.000000		

Para el análisis de los resultados nos concentramos en el valor del Coeficiente de Correlación R^2 y el valor de Durbin – Watson, los cuales se pueden observar en la tabla (5-15) y cuyos respectivos valores son 0.908663 y 1.458439 , lo que indica que el modelo presenta buena calidad para replicar los resultados, y expresa el 90.86 % de la varianza de la variable dependiente Ozono (O_3) explicada por las variables independientes que en este caso con la velocidad del viento y la variable de material particulado $PM_{2.5}$ y el valor de Durbin – Watson nos dice que los residuos no están correlacionados.

d) Validación del modelo.

I De acuerdo con la tabla (5-15) $R^2 = 0.90866$

II De la tabla (5-15) Durbin-Watson stat= 1.458439

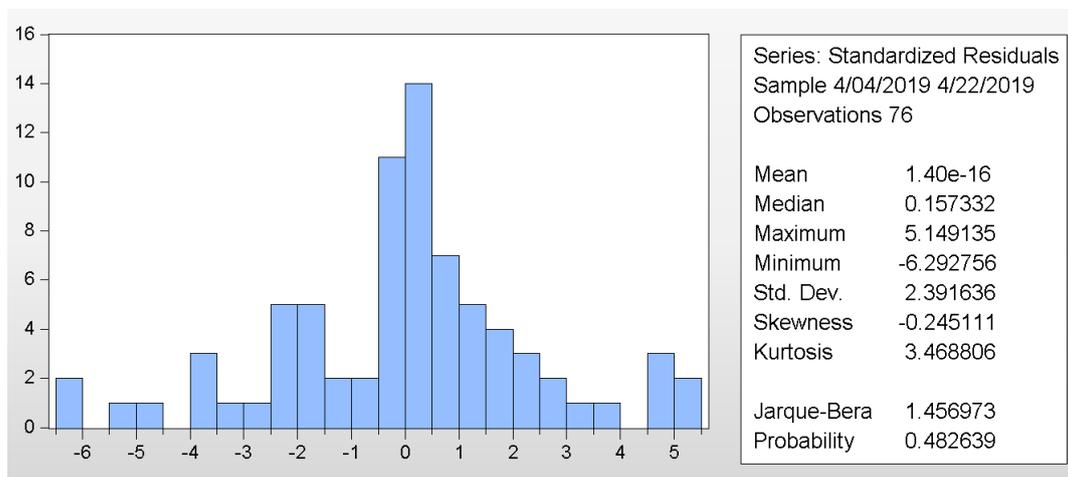
III Normalidad. Para la normalidad normalidad se utiliza el test de Jarque Bera.

Contraste de hipótesis

H_0 : Los residuales son normales.

H_1 : Los residuales no son normales.

Figura 5-7: Prueba de normalidad



Se observa en la figura (5-7) que el p -valor del contraste de normalidad de Jarque Bera es superior a 0.05; lo que indica residuos normales con una confianza del 95 %.

IV Para probar si los efectos fijos de las estaciones y de tiempo pueden o no considerarse iguales se utiliza el test de máxima verosimilitud para la redundancia. En la tabla (5-16) se observan p -valores menores que 0.05 , lo que nos lleva a afirmar que los efectos fijos de las estaciones y de tiempo son diferentes con un 95 % de confianza.

Tabla 5-16: Add your caption here.

Redundant Fixed Effects Tests
Test cross-section and period fixed effects

Effects Test	Statistic	d.f.	Prob.
Cross-section F	39.624344	(3,52)	0.0000
Cross-section Chi-square	90.415456	3	0.0000
Period F	8.519830	(18,52)	0.0000
Period Chi-square	104.386450	18	0.0000
Cross-Section/Period F	16.873596	(21,52)	0.0000
Cross-Section/Period Chi-square	156.252967	21	0.0000

V Prueba de dependencia de los residuales Las pruebas de dependencia de sección transversal se muestran en la tabla (5-17)

Tabla 5-17: Prueba dependencia residuales.

Residual Cross-Section Dependence Test
Null hypothesis: No cross-section dependence (correlation) in residuals

Test	Statistic	d.f.	Prob.
Breusch-Pagan LM	37.33479	6	0.0000
Pesaran scaled LM	9.045574		0.0000
Bias-corrected scaled LM	8.934463		0.0000
Pesaran CD	-3.103170		0.0019

La primera línea contiene los resultados de la prueba Breusch-Pagan LM. En este caso, el valor del estadístico de prueba, 37.33479, y rechazamos hipótesis nula de no correlación en los niveles de significación convencionales. Las siguientes líneas presentan resultados para las dos pruebas de Breusch-Pagan escaladas y para Pesaran las cuales rechazan fuertemente la hipótesis nula a un 5%

VI Heterocedasticidad

Tabla 5-18: Prueba heterocedasticidad

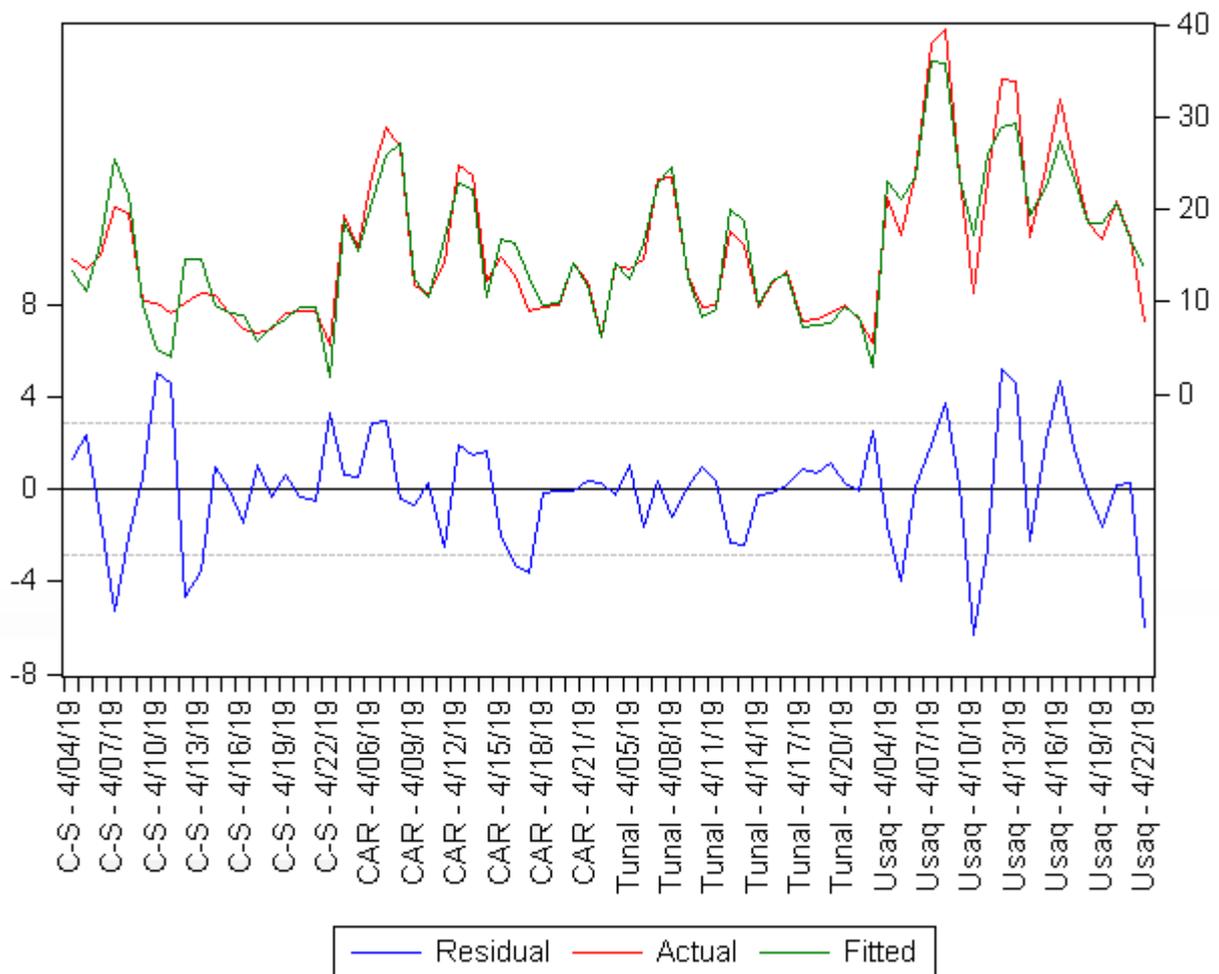
Test for Equality of Variances of RESID
Categorized by values of RESID

Method	df	Value	Probability
Bartlett	4	3.463855	0.4834
Levene	(4, 71)	1.841642	0.1304
Brown-Forsythe	(4, 71)	1.752136	0.1482

La tabla (5-18) muestra los resultados, observándose que se acepta la igualdad de varianzas residual en las distintas secciones cruzadas dado que los p -valores en cada una de las pruebas con mayores que 0.05. No existe entonces heteroscedasticidad

e) Gráfico de estimación

Figura 5-8: Gráfico de estimación



f) Efectos individuos - Efectos tiempo

Tabla 5-19: Cross-Section Fixed effects O₃

INDIVIDUO	Effect
C-S	0.557482
CAR	-1.308303
Tunal	-4.614369
Usaq	5.365191

En la tabla (5-19) se evidencia que en los sectores de Usaquén, y Carvajal-Sevillana hay efecto positivo y mientras que en el Tunal y Centro de alto rendimiento hay un efecto negativo sobre la variable del ozono.

Tabla 5-20: Period Fixed Effects.

Fecha	Effect
2019 - 04 - 04	-0.756336
2019 - 04 - 05	1.285092
2019 - 04 - 06	2.288646
2019 - 04 - 07	11.05774
2019 - 04 - 08	15.54993
2019 - 04 - 09	2.379624
2019 - 04 - 10	-1.014288
2019 - 04 - 11	3.812599
2019 - 04 - 12	7.582769
2019 - 04 - 13	6.529668
2019 - 04 - 14	-3.259317
2019 - 04 - 15	-1.387884
2019 - 04 - 16	-0.100270
2019 - 04 - 17	-2.019491
2019 - 04 - 18	-7.705698
2019 - 04 - 19	-9.921820
2019 - 04 - 20	-6.328531
2019 - 04 - 21	-8.599566
2019 - 04 - 22	-9.392865

En la tabla (5-20) se observa que los días de mayor y menor aporte fueron el 2019-04-19 y el 2019-04-08 respectivamente

g) Ecuación para estimar los niveles de Ozono O_3

A continuación se presenta la tabla la ecuación para la estimación de la variable de respuesta, es decir para el Ozono

$$O_3 = 22.51066 - 0.418676PM_{2.5} + 2.954323velviento + 0.557482d_1 - 1.308303d_2 - 4.614369d_3 + 5.365191d_4 + -0.756336A_4 + 1.285092A_5 + 2.288646A_6 + 11.05774A_7 + 15.54993A_8 + 2.379624A_9 - 1.014288A_{10} + 3.812599A_{11} + 7.582769A_{12} + 6.529668A_{13} - 3.259317A_{14} - 1.387884A_{15} - 0.100270A_{16} - 2.019491A_{17} - 7.705698A_{18} - 9.921820A_{19} - 6.328531A_{20} - 8.599566A_{21} - 9.392865A_{22}$$

($d_i = 1$ para observaciones de la estación i y vale $d_i = 0$ en caso contrario, de la misma forma $A_t = 1$ para observaciones del día t y $A_t = 0$ en caso contrario)

5.3. Análisis Multivariado

En esta sección se realizará una Análisis de Correspondencia múltiple para las variables PM₁₀, PM_{2,5}, SO₂, NO₂, O₃ CO, Precipitación, Dirección del viento, velocidad del viento y precipitación cada una de estas debidamente categorizadas

El Ministerio de Ambiente y Desarrollo Sostenible y la Secretaría Distrital de Ambiente de Bogotá, acaban de lanzar el Índice Bogotano de Calidad de Aire – IBOCA. El IBOCA es un indicador Multipropósito Adimensional, que incluye para su cálculo las siguientes variables de calidad del aire: material particulado menor a 10 μm (PM₁₀), material particulado menor a 2.5 μm (PM_{2,5}), ozono (O₃), monóxido de carbono (CO), dióxido de azufre (SO₂) y dióxido de nitrógeno (NO₂).

En la figura (5-9) se muestran los rangos de concentración de contaminantes del aire y los diferentes atributos del IBOCA:

Figura 5-9: Índice de la calidad del aire



ÍNDICE BOGOTANO DE CALIDAD DEL AIRE

Atributos del IBOCA				Rangos de concentración y tiempo de exposición para cada contaminante ¹					
Rangos numéricos	Color	Estado de calidad del aire	Estado de actuación y respuesta ²	PM10, 24h (µg/m³)	PM2.5, 24h (µg/m³)	O ₃ , 8h (µg/m³) [ppb]	CO, 8h (µg/m³) [ppm]	SO ₂ , 1h (µg/m³) [ppb]	NO ₂ , 1h (µg/m³) [ppb]
0 - 10	Azul claro	Favorable	Prevención	(0-54)	(0-12)	(0-116) [0-59]	(0-5038) [0.0-4.4]	(0-93) [0-35]	(0-100)
10,1 - 20	Verde	Moderada	Prevención	(55-154)	(12.1-35.4)	(117-148) [60-75]	(5039-10762) [4.5-9.4]	(94-198) [36-75]	(101-188)
20,1 - 30	Amarillo	Regular	Alerta Amarilla	(155-254)	(35.5-55.4)	(149-187) [76-95]	(10763-14197) [9.5-12.4]	(199-486) [76-185]	(189-677) [101-360]
30,1 - 40	Naranja	Mala	Alerta Naranja	(255-354)	(55.5-150.4)	(188-226) [96-115]	(14198-17631) [12.5-15.4]	(487-797) [186-304]	(678-1221) [361-649]
40,1 - 60	Rojo ³	Muy Mala	Alerta Roja ³	(355-424)	(150.5-250.4)	(227-734) [116-374]	(17632-34805) [15.5-30.4]	(798-1583) [305-604]	(1221-2349) [650-1249]
60,1 - 100 ⁴	Morado	Peligrosa	Emergencia	(425-604)	(250.5-500.4)	(734-938) [374-938]	(34806-57703) [30.5-50.4]	(1584-2630) [605-1004]	(2350-3853) [1250-2049]

Tomada: <http://www.icde.org.co/noticias/Indice-Bogotano-de-Calidad-de-Aire-IBOCA>

Antes de realizar El MCA se muestran en las figuras

Figura 5-10: PM_{2.5} por individuo

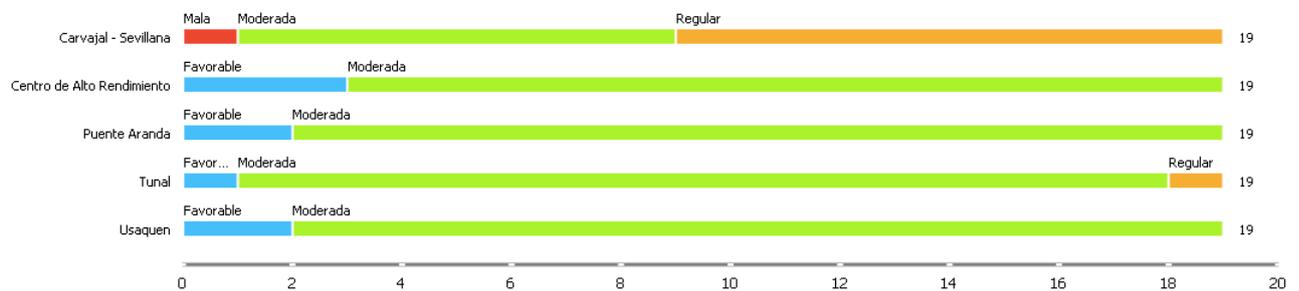


Figura 5-11: PM₁₀ por individuo

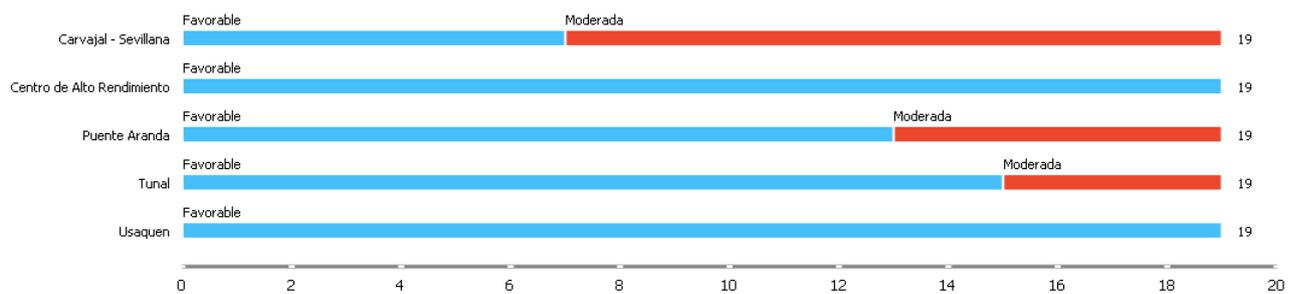
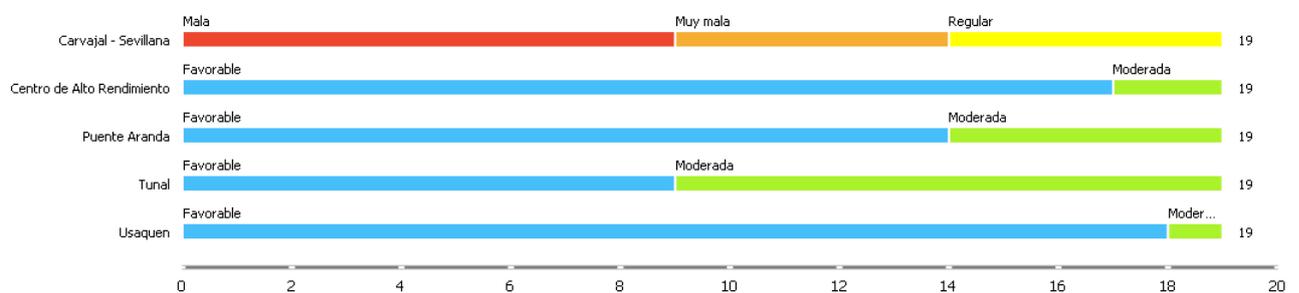
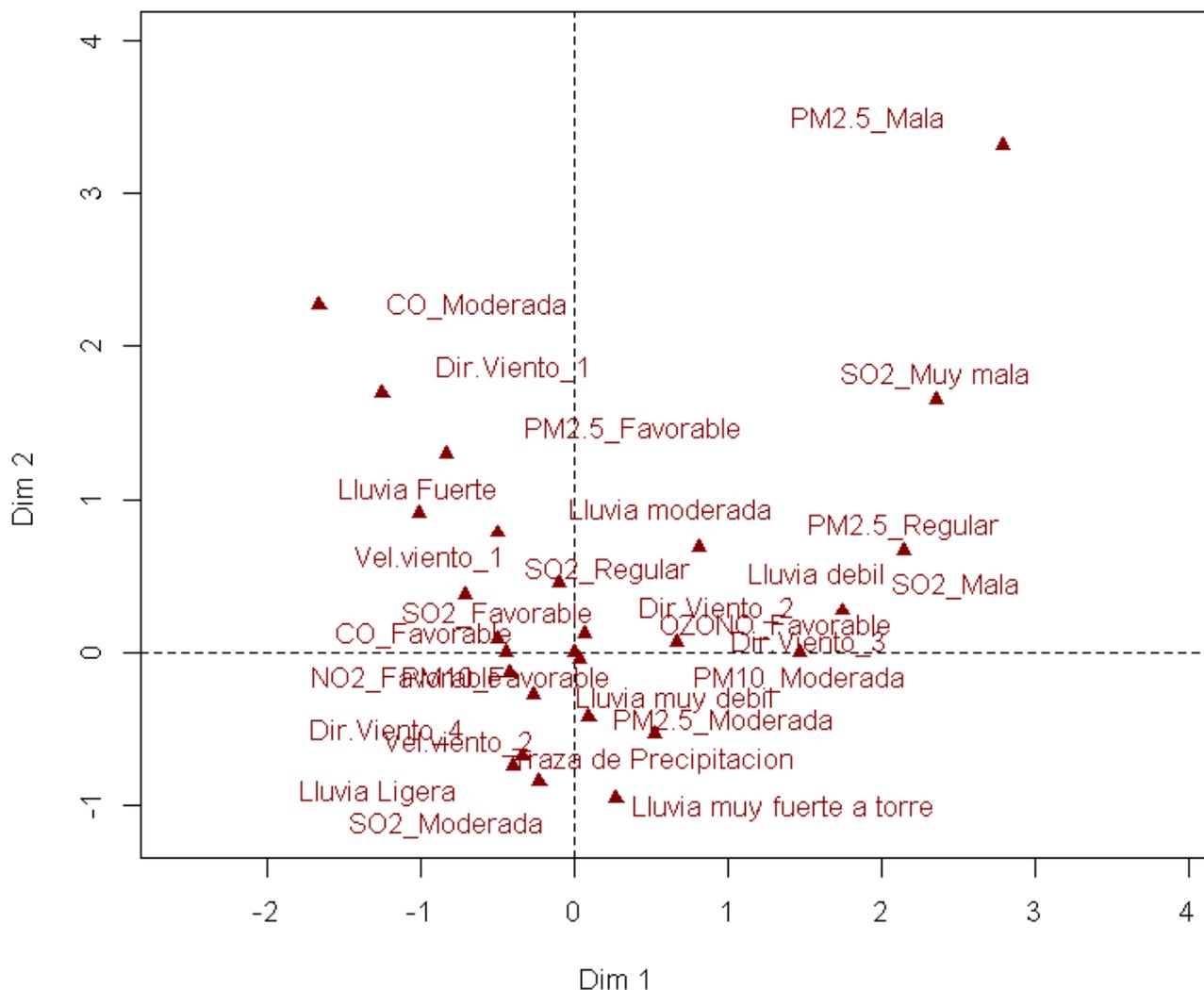


Figura 5-12: SO₂ por individuo



Al realizar el MCA entre las diferentes categorías de cada una de las variables se obtiene como plano principal el mostrado en la figura (5-13).

Figura 5-13: MCA Categorías

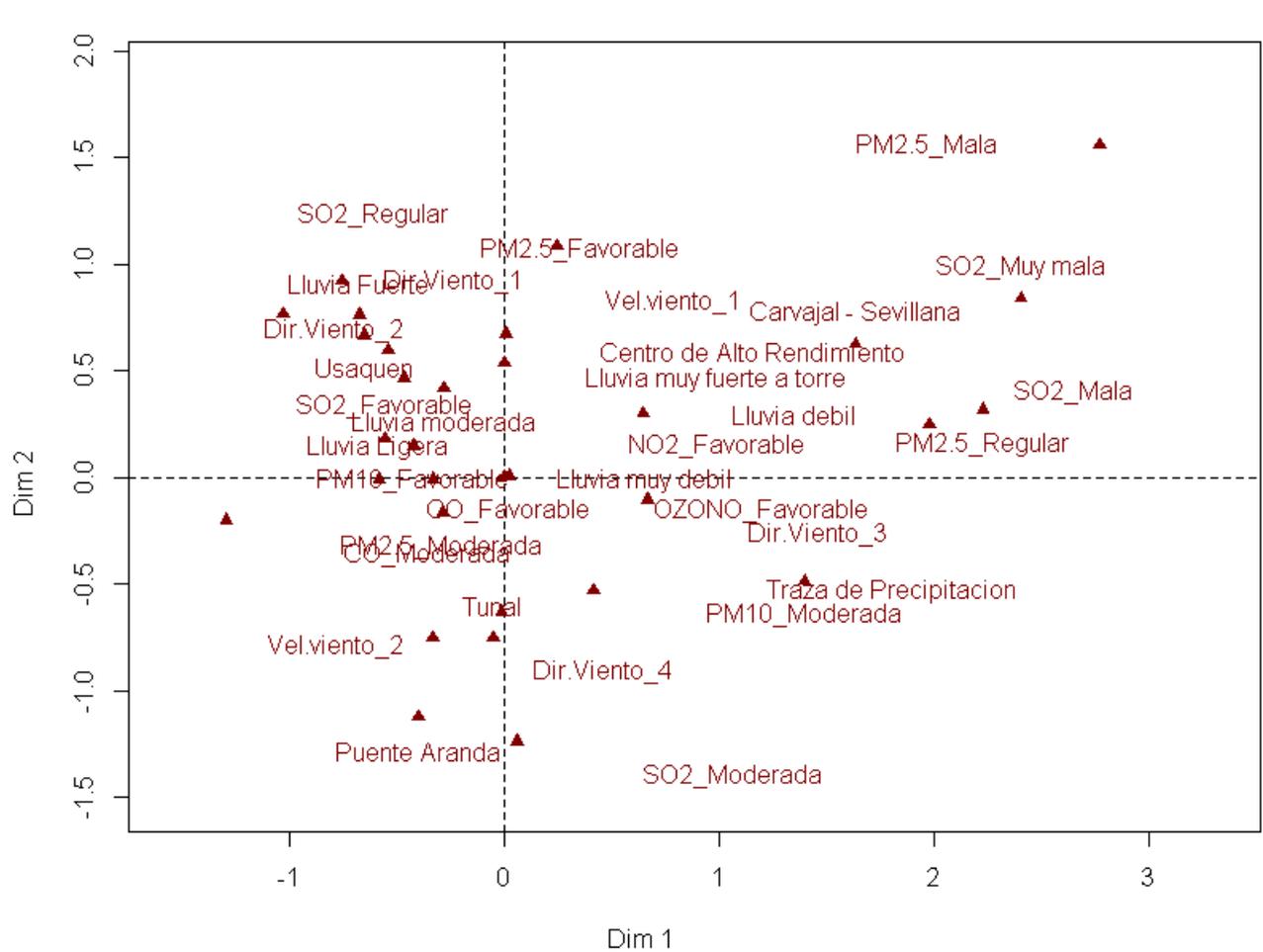


En la gráfica (5-13) se observa que que hay unos puntos que están bien representados en la primera dimensión como lo son los correspondientes al material particulado de la variable $PM_{2.5}$ de su categoría de igual manera se observa que la variables SO_2 en su categoría Muy mala, $PM_{2.5}$ Regular y SO_2 Mala están bien representadas en esta dimensión

A continuación se muestra la gráfica del MCA, pero en este caso se incluyen los individuos en este caso las 4 zonas de Bogotá donde se esta realizando el estudio, como los son

Carvajal-Sevillana, Tunal, Centro de Alto rendimiento y Usaquén obteniéndose como resultado la figura (5-14)

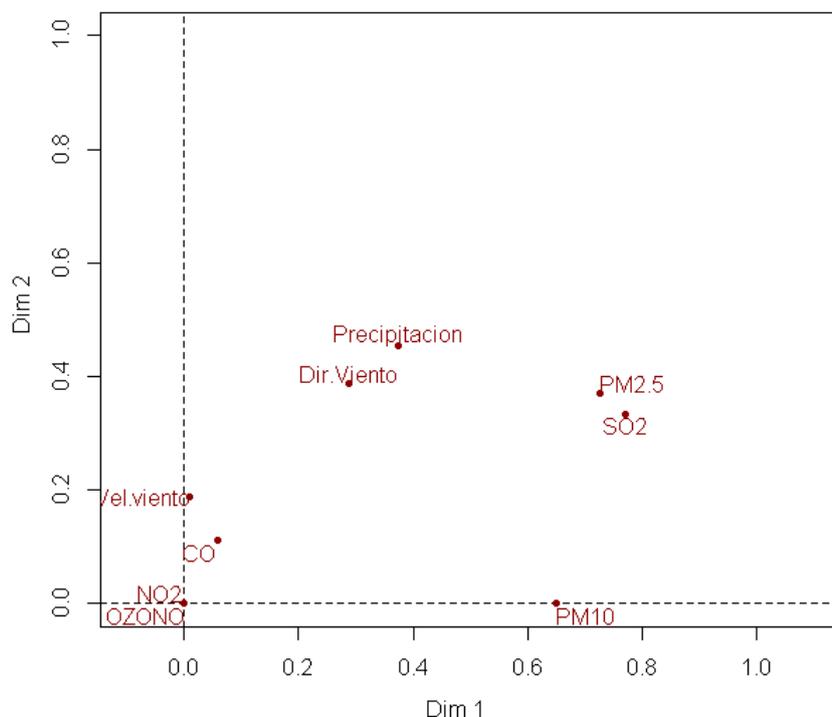
Figura 5-14: MCA categorías con individuos



En la gráfica (5-14) se observa hay unas unas categorías que están bien representados en la primera dimensión como lo son los correspondientes a PM_{2.5}Mala SO₂ Muy mala, PM_{2.5}Regular, SO₂ Mala, y muy cerca se observa que cerca a ellas se encuentra Carvajal-Sevillana, lo indica que hay correlación entre estas categorías y esta zona.

En la figura (5-15) se muestra el MCA solo para las variables

Figura 5-15: MCA variables



En la figura (5-15) se observa que existe una correlación entre las variables $PM_{2.5}$ y SO_2 , de igual manera se puede afirmar que estas variables están quedando representadas en la dimensión 1 junto con el material particulado PM_{10} . En la segunda dimensión hay aportes de las variables: Precipitación, Dirección del viento, velocidad del viento y $PM_{2.5}$ y SO_2 . Mientras que las variables Ozono (O_3) y dióxido de nitrógeno (NO_2) no están representadas en estas dimensiones.

Lo anterior puede ser justificado ya que de acuerdo con estudios (??) muestran que los principales contaminantes del aire es el material particulado pues superan frecuentemente los valores por la regulación ambiental, y se afirma El aire de Bogotá no presenta un problema de contaminación del aire para NO_2 , CO y O_3 para estos contaminantes registradas por la red de monitoreo de la calidad del aire de Bogotá suelen ser inferiores a los límites establecidos por la regulación ambiental excepto

En la figura (5-16) se puede corroborar el aporte que hacen cada una de las categorías de las variables en las primeras 5 dimensiones.

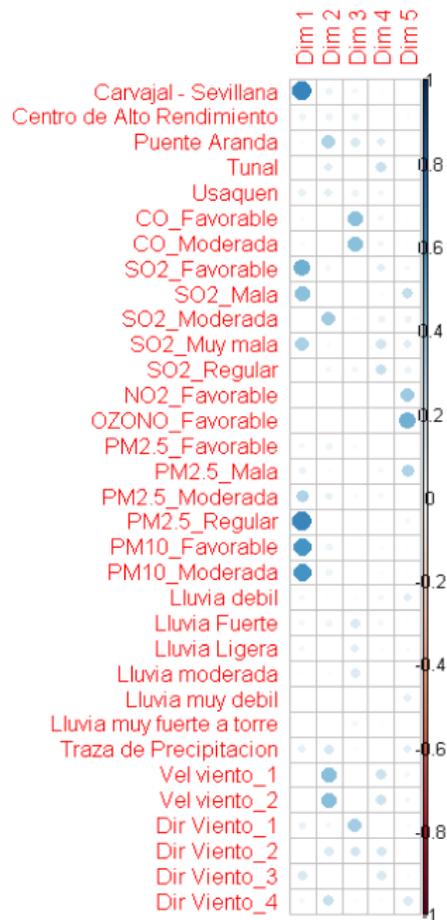


Figura 5-16: mca cos²

Capítulo 6

Conclusiones y Recomendaciones

- Analizando los efectos estimados, se encontró que la zona de Carvajal-Sevillana es la que más aporta sobre la variable SO_2 y el $\text{PM}_{2.5}$ indicando que en esta zona de la ciudad los niveles de estos elementos son altos y se puede determinar que presentan un mayor nivel de contaminación con respecto a las demás zonas de la ciudad analizadas en este trabajo, y en según el análisis por individuo se evidencia que las demás zonas analizadas se encuentran favorables y moderadas de acuerdo con el índice de calidad del aire para Bogotá y la variable meteorológica que tiene influencia es el viento esto por la relación existente entre las concentraciones de emisores contaminantes y las corrientes de viento.
- El modelo de datos panel que determina la concentración de SO_2 incluye las variables $\text{PM}_{2.5}$, velocidad y dirección del viento; y para el O_3 incluye las variables $\text{PM}_{2.5}$, velocidad del viento para las cuatro zonas de Bogotá analizadas.
- Para el gas contaminante atmosférico SO_2 en las cuatro zonas de Bogotá analizadas se encontró que la zona que más aporta a la concentración de este gas es Carvajal - Sevillana, por ser una zona industria l donde están localizadas grandes fabricas, un río en condiciones inadecuadas, el matadero más grande de la zona y dos de las vias de mayor congestión, recorridas por vehículos de trafico pesado y la zona con menor aporte fue Usaquen; para los datos analizados la fecha que apporto en mayor medida fue el 19 de abril de 2019 viernes y el que brindo menor aporte fue 08 de abril de 2019.

Bibliografía

- Abel, G., y Flórez, R. (2018). Conpes 3943 Política para el mejoramiento de la calidad del aire.
- Alcaldía Mayor de Bogotá D.C. (s.f.). *Ambiente Bogotá - Secretaria Distrital de Ambiente*. Descargado de <http://www.ambientebogota.gov.co/web/sda/inicio>
- Baronio, A., y Vianco, A. (2014). *Datos de Panel Guía para el uso de Eviews* (Inf. Téc.). Universidad Nacional de Río Cuarto.
- Galvis, B., y Rojas, N. (2006). Relación entre PM_{2.5} y PM₁₀ en la Ciudad de Bogotá. , 336–353.
- Mayorga, M., y Evelyn Muñoz, M. (2000). *LA TÉCNICA DE DATOS DE PANEL UNA GUÍA PARA SU USO E INTERPRETACIÓN* (Inf. Téc.). Descargado de <https://www.mendeley.com/viewer/?fileId=532301ea-f17d-3ea7-1584-c66e4f9cf65e&documentId=bc96ae2c-ef0b-3998-a162-7aa2dce4fdcb>
- Nadal, I. S. (2006). Fundamentos de meteorología. *Univ Santiago de Compostela.*, 6.
- O.M.S. (2018). *Organización Mundial de la Salud*. Obtenido de. Descargado de [http://www.who.int/es/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](http://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- OPS OMS — *Calidad del aire — Productos científicos y técnicos*. (s.f.). Descargado de https://www.paho.org/hq/index.php?option=com_topics&view=rdmore&cid=8034&item=calidad-aire-contaminacion&type=scientific_technical&Itemid=40799&lang=es
- Perez, C. (s.f.). *Econometría avanzada. Técnicas y herramientas. Econometría de los datos de panel* (4.^a ed.; Cengage Learning Editores, Ed.).
- WHO — Ambient and household air pollution and health. (2018). *WHO*. Descargado de <https://www.who.int/airpollution/data/en/>
- Wooldridge, J. M., Enfoque Moderno, U., Del Carmen, M., Hano, E., Érika, R., Jasso, M., ... Benavides, D. R. (2009). *Introducción a la econometría Traducción Revisión técnica* (Inf. Téc.).